Medicine
OPEN

Comparison of prediction accuracies between two mathematical models for the assessment of COVID-19 damage at the early stage and throughout 2020

Hua-Ying Chuang, MD[a,b], Tsair-Wei Chien, MBA[b,c] iD, Willy Chou, MD[d,e], Chen-Yu Wang, MS[f], Kang-Ting Tsai, MD[g,h]*

Abstract

Background: The negative impacts of COVID-19 (ImpactCOVID) on public health are commonly assessed using the cumulative numbers of confirmed cases (CNCCs). However, whether different mathematical models yield disparate results based on varying time frames remains unclear. This study aimed to compare the differences in prediction accuracy between 2 proposed COVID-19 models, develop an angle index that can be objectively used to evaluate ImpactCOVID, compare the differences in angle indexes across countries/regions worldwide, and examine the difference in determining the inflection point (IP) on the CNCCs between the 2 models.

Methods: Data were downloaded from the GitHub website. Two mathematical models were examined in 2 time-frame scenarios during the COVID-19 pandemic (the early 20-day stage and the entire year of 2020). Angle index was determined by the ratio (=CNCCs at IP÷IP days). The R^2 model and mean absolute percentage error (MAPE) were used to evaluate the model's prediction accuracy in the 2 time-frame scenarios. Comparisons were made using 3 visualizations: line-chart plots, choropleth maps, and forest plots.

Results: Exponential growth (EXPO) and item response theory (IRT) models had identical prediction power at the earlier outbreak stage. The IRT model had a higher model R^2 and smaller MAPE than the EXPO model in 2020. Hubei Province in China had the highest angle index at the early stage, and India, California (US), and the United Kingdom had the highest angle indexes in 2020. The IRT model was superior to the EXPO model in determining the IP on an Ogive curve.

Conclusion: Both proposed models can be used to measure ImpactCOVID. However, the IRT model (superior to EXPO in the long-term and Ogive-type data) is recommended for epidemiologists and policymakers to measure ImpactCOVID in the future.

Abbreviations: CNCC = cumulative number of confirmed cases, EXPO = exponential growth, ImpactCOVID = impacts of the COVID-19 outbreak, IRT = item response theory, LID = the length of infected days effectively under control, MA = model accuracy, MAPE = mean absolute percentage error, PP = prediction power, QE = quadratic equation.

Keywords: COVID-19, exponential growth model, inflection point, IRT model, item response theory, MAPE, Ogive curve, Ogive curve, R-square

1. Introduction

In the field of epidemiology, an "outbreak" refers to a sudden increase in occurrences of a disease at a particular time and place. For comparison, a "pandemic" is defined as a near-global disease outbreak in which multiple countries across the world are infected.[1] The cumulative numbers of confirmed cases (CNCCs) are a commonly used indicator to assess the negative impacts of the COVID-19 outbreak (ImpactCOVID for short).[2–5] However, this practice might be questionable because CNCCs do not involve the length of infected days to control COVID-19 (denoted by LID). This concept is reflected in the inflection point (IP) days on the CNCC curve, wherein the curvature changes its sign from an increasing concave (concave downward) to a

---

decreasing convex (concave upward) shape or vice versa.[1] Hence, combining CNCCs and IP-based LID for measuring ImpactCOVID is reasonable and necessary.

1.1. Literature review

1.1.1. Using IP and CNCCs to assess ImpactCOVID. If CNCCs and IP-based LID are considered, then ImpactCOVID (= CNCCs ÷ LID) is similar to the journal impact factor (i.e., IF = citations ÷ publications). Many metrics, such as the h- and x-indexes, have been proposed to improve the drawbacks of impact factors.[6,7] The h-index is determined by the maximum square that fits under an author's citation curve when the numbers of citations are plotted in decreasing order, and the x-index is determined by the maximum area rectangle that fits under the curve.[7] All excessive citations and publications are excluded from the metric computation.

The IPcase index, which uses the rectangular area ("core area") multiplied by the IP days and the CNCCs, has been proposed to identify ImpactCOVID.[1,4] However, this index has 2 drawbacks: it has an identical value but a distinct meaning (e.g., IPcase index = CNCC × IP = height × width = 100 = 4 × 25 = 25 × 4 = 10 × 10), and it does not denote the momentum of a sudden ImpactCOVID. Thus, the improved angle index is defined as follows:

$$\text{angle index } = \theta = \text{Degrees}(\text{Atan}(\frac{\Delta\text{CNCC}_k}{(\Delta\text{IP}_k)})), \quad (1)$$

where Degrees() and Atan() are derived from the functions in Microsoft Excel. For instance, ΔIP = 7 days, ΔCNCC = 27,100 − 11,177 = 15,934, ratio = 15,934/(7−1)= 2563, θ = Degrees(Atan(2653.8)) = 89.97. The angle index ranges from 0 to 90, wherein a high θ value means a great negative impact ("severely hit") by COVID-19 in a given country or region. The premise is to determine the IP days (similar to the LID defined in the previous section) before calculating the angle index for a specific country/region.

1.1.2. Using mathematical models to determine the IP. In general, the mean number of confirmed cases across varying periods (number of days) yields significantly different IP days, even though the daily number of confirmed cases (computed based on the previous 7 days) can be applied to estimate the observed IP days.[8,9] In a similar situation, using a mathematical model to determine the expected IP days is more objective than traditionally observing the IP days on the mean number of confirmed cases; the latter is typically but irrationally employed in practice.[2]

Although many mathematical models[2,10–21] have been proposed to predict the number of COVID-19 cases, none (except the one using item response theory [IRT][2,20,21]) have been applied to determine the IP days on the CNCC curve. Nonetheless, the differences in model accuracy (MA) and prediction power (PP) should be compared. MA and PP can be measured by the R^2 model and mean absolute percentage error (MAPE), respectively (see the Methods section for details on R^2 and MAPE).

Therefore, we applied MA and PP as indicators in the 2 COVID-19 models (i.e., exponential growth (EXPO) and IRT)[2,18] and used the angle index to measure ImpactCOVID for each country/region.

1.2. Main Goals

This study aimed to compare the differences in MA and PP between the 2 proposed COVID-19 models, develop an angle index that can be objectively used to evaluate ImpactCOVID, compare the differences in angle indexes across countries/regions worldwide, examine the difference in determining the
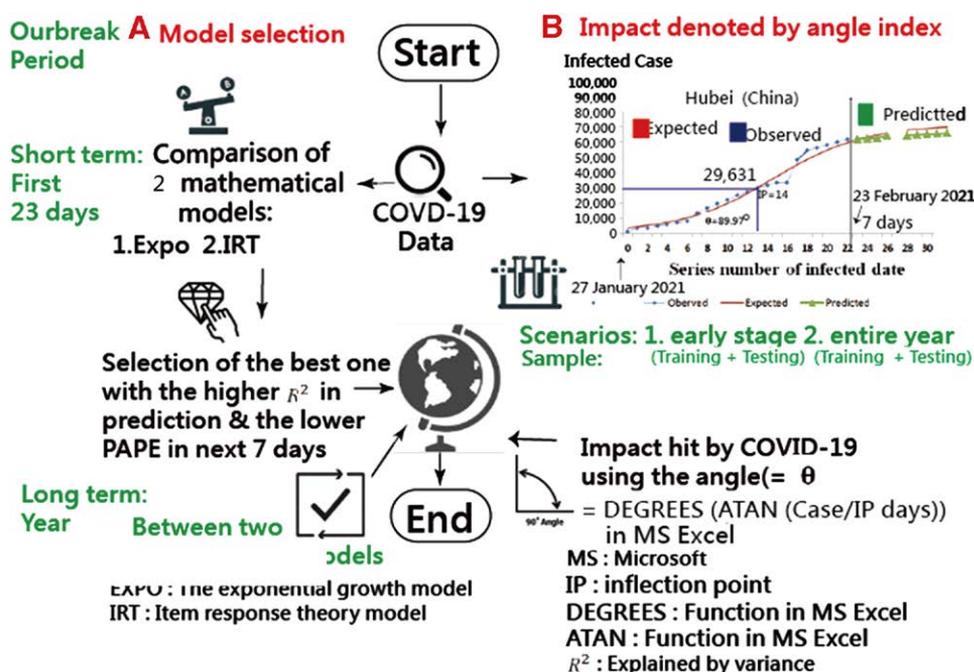
Figure 1. Study flowchart.

Indicator	Formula
AUC	$=CC*CC$
Residual	$=\sum(i = 1\ \text{to}\ n)[(Oi - Ei)^{\wedge}2]$
AIC	$=-2\ \ln(L) + 2\ k = n*\ln(SSE/n)+2k$
BIC	$=2\ \ln(L) + \ln(n)*k$
Brier score	$=\sum(i = 1\ \text{to}\ n)[(Oi - Ei)^{\wedge}2]/n$

SSE = $\sum(i = 1\ \text{to}\ n)[(Oi - Ei)^{\wedge}2]$ =the sum of squares due to error = Residual; n = observed counts; CC = correlation coefficient; Ln(L) = log(maximum likelihood); Oi = the ith observed value; Ei = the ith expected value; k = the number of model parameters; AUC = area under the curve = determination coefficient = CC*CC.

IP on the CNCC curve between the 2 mathematical models, and compare goodness-of-fit measures for the 2 models.[2,18]

2. Materials

2.1. Data source

COVID-19 data have been obtained from GitHub[22] for each country/region[23] (see Supplemental Digital Contents 1, http://links.lww.com/MD/G983 and 2, http://links.lww.com/MD/G984). The downloaded data are publicly available on the website.[22] It was not necessary to obtain ethical approval for this study since all data were obtained from the GitHub website.

2.2. Introducing the mathematical models

2.2.1. The EXPO model. In the study of natural science, nonlinear regression and iterative methods are commonly employed.[24,25] As part of the COVID-19 prediction model,

EXPO has been proposed,[18] which is based on the daily growth rate (GR) of confirmed cases using Eq. 2:

$$GR_{n-1} = \frac{case_n - case_{n-1}}{case_{n-1}}, \qquad (2)$$

$$GR_{t-1} = ae^{-\beta t}, \qquad (3)$$

where $case_n$ and $case_{n-1}$ are the daily number of confirmed cases on days n and $n - 1$, respectively. The GR can be modeled in Eq. 3, where a is a constant representing the growth rate at $t = 0$, β is an attenuation coefficient that indicates the effectiveness of government isolation and quarantine, and t represents the evolution of the epidemic.[18] Based on Eq. 2, nonlinear regression and iterative methods can be constructed to predict the CNCCs using Eq. 4:

$$CNCC_k = CNCC_0 \times \prod_{i=1}^{k}(1 + ae^{-\beta t}). \qquad (4)$$

where CNCC_0 and CNCC_k are the CNCCs at $t = 0$ and the expected CNCC at $t = k$, respectively (see Supplemental Digital Content 1, http://links.lww.com/MD/G983).

2.2.2. The IRT Model. The item response model (IRT) was proposed in 2021[2,20,21] using Eq. 5:

$$P(\theta) = \frac{1}{1 + e^{-1.7*a(\theta - b)}} = \frac{e^{1.7*a(\theta - b)}}{1 + e^{1.7*a(\theta - b)}}, \qquad (5)$$

The *a* and *b* parameters represent the discrimination (i.e., a slope from 0 to 4) and the difficulty (i.e., a value from –5 to 5, with the left value indicating that the outbreak started earlier and the right value indicating that it lasted longer).[2,20,21] The number of infected days denoted by θ is standardized to range between –5 and 5 (see Supplemental Digital Content 2, http://links.lww.com/MD/G984).
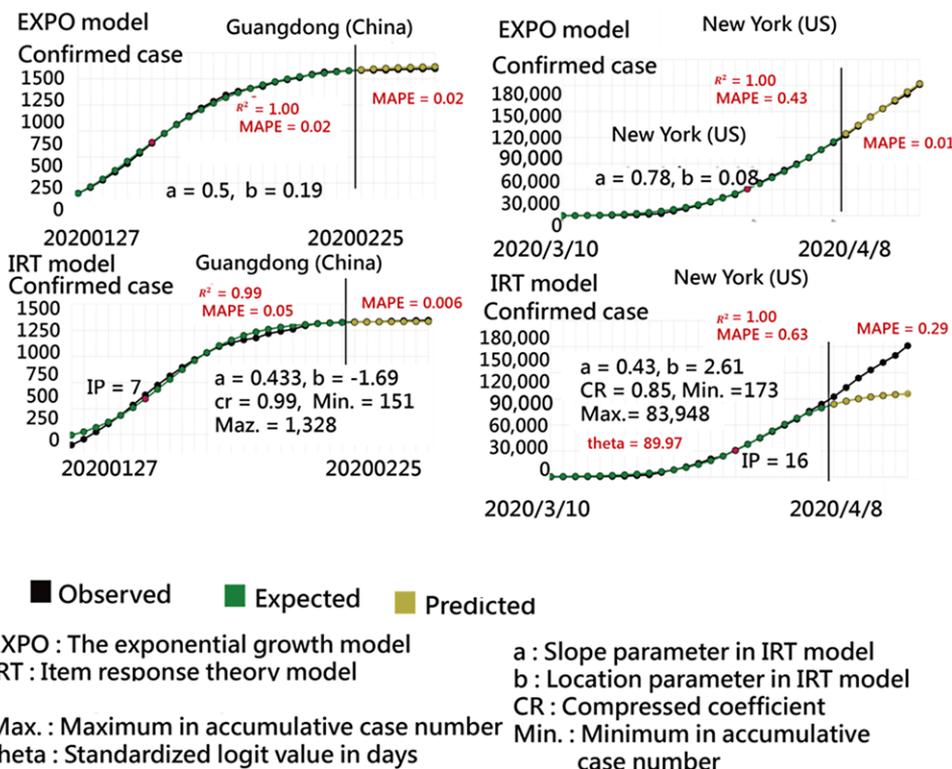
Figure 2. Comparison of accuracy among the proposed models.

2.2.3. Model parameter estimations. The model parameters can be estimated by minimizing the residual using the Microsoft add-in tool.[2,20] Details are provided in Supplemental Digital Contents 1, http://links.lww.com/MD/G983 and 2, http://links. lww.com/MD/G984 (e.g., executing the procedure, SolverSolve UserFinish: = True, ShowRef: = "ShowTrial," using visual basic for applications, VBA).[2,20,21,23]

2.2.4. Analyzing model accuracy and prediction power. Both R^2 and MAPE statistics were applied to evaluate the model accuracy (MA) and prediction power (PP) in the 2 time-frame scenarios (see the training and testing samples shown in Fig. 1). We applied the model parameters calibrated in the training sample to predict the CNCCs in the testing sample by observing the 2 measures of R^2 and MAPE. The R^2 and MAPE are defined in Eqs. 7 and 8:

$$R^2 = 1 - \frac{\text{model residual}}{\sum_{i=1}^{n} (O_i - \text{Ohat})^2} \quad (7)$$

$$\sum_{i=1}^{n} (O_i - E_i)^2 \quad (8)$$

where O_i denotes the observed CNCCs and O-hat is the mean CNCC in a given country/region. The model residual is computed by Eq. 8. E_i represents the expected CNCCs.

$$\text{MAPE} = \frac{1}{n} \times \sum_{i=1}^{n} |\frac{(O_i - E_i)}{O_i}| \quad (9)$$

where O_i is the observed CNCC and E_i is the predicted CNCC. The absolute value in Eq. 8 is summed across all predicted points in days and divided by the number of fitted points n.

2.3. Three tasks using forest plots in model comparisons

The forest plot[26] was used to examine the difference in R^2 and MAPE. The first task was to compare the effects of R^2 and MAPE in the 2 models at the early outbreak stage. The second task was to observe their difference in R^2 and MAPE throughout 2020. The 3 tasks were to compare their IP days on vaccination data from January 1 to November 6 in 2021. The angle index defined in Eq. 1 was used to measure ImpactCOVID (see the example in the Introduction).
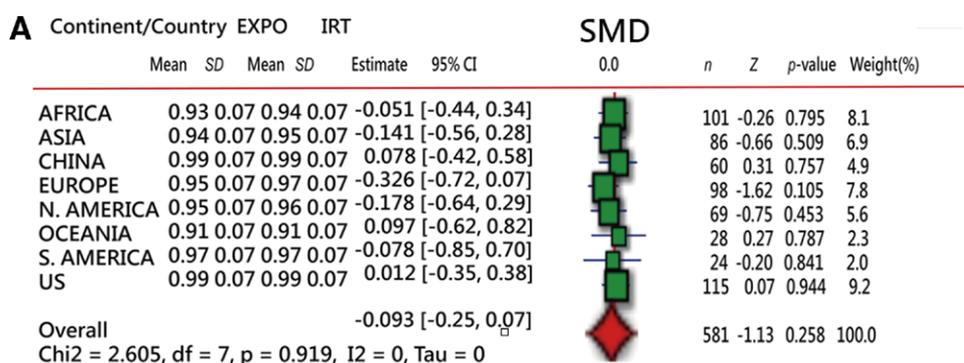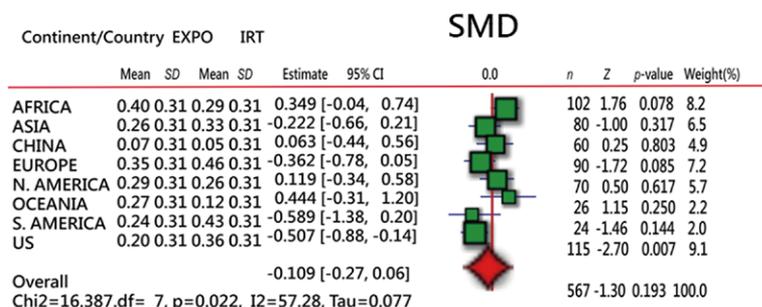
Figure 3. Comparison of R^2 and MAPE across continents, China, and the US based on the 3 models. MAPE = mean absolute percentage error.
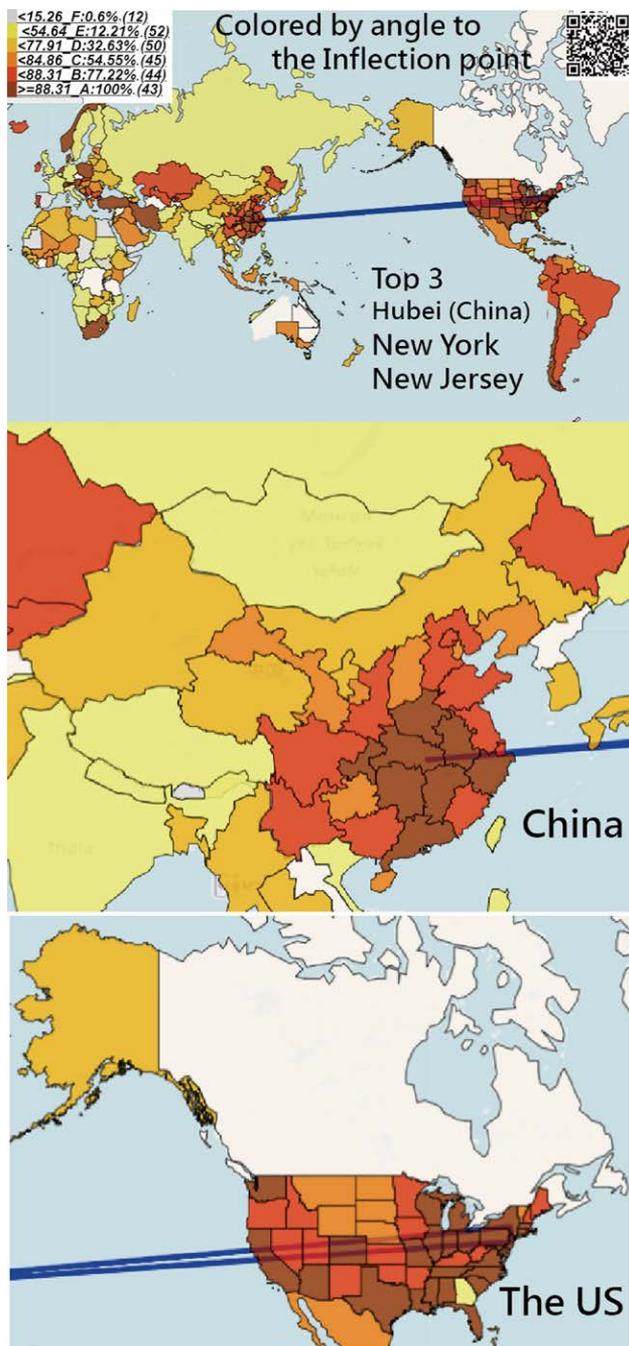
Figure 4. ImpactCOVID based on the angle index at the earlier stage. ImpactCOVID = impacts of the COVID-19 outbreak.

We used choropleth maps[27] to present the angle indexes for countries/regions at an early stage and throughout 2020. The darker color in countries/regions indicates more severe impacts as measured by ImpactCOVID. On clicking a colored region, line plots that are used to predict the future CNCCs are generated.

2.4. Task 4: goodness-of-fit measures for the 2 models

Using the goodness-of-fit measures shown in Table 1, 5 indicators were compared between the 2 models.[2,18] The area under the curve (AUC) is defined by the regression R^2 (i.e., the determination coefficient computed by plotting the Pearson correlation coefficients between the 2 variables of the last 7-day predicted

and observed confirmed cases in the 2 scenarios of the 20-day early outbreak stage and throughout 2020) and drawn on a scatter plot with the 95% control lines[28] for countries/regions based on the IRT and EXPO mode ls.

The residuals on the two 7-element variables were computed by Eq. 8. The other 3, including the Akaike information criterion,[29] the Bayesian information criterion,[30] and the Brier score,[31] are compared. A smaller value means better modeling of the COVID-19 pandemic.

2.5. Statistical tools and data analysis

The mean and standard deviation (SD) were extracted to compare the standardized mean difference (SMD) in the forest plot. A significance level of type I error was set at 0.05.

Visual displays of the forest plot and choropleth map illustrate the comparison between MA and PP. The angle indexes were plotted online on Google Maps. The parameter estimation was executed in Microsoft Excel[23] (Supplemental Digital Contents 1, http://links.lww.com/MD/G983 and 2, http://links.lww.com/MD/G984). The study flowchart is shown in Figure 1.

3. Results

3.1. Task 1

3.1.1. Model comparisons at the early outbreak stage. An example of Guangdong in China and New York in the United States was used to illustrate the comparison. The results show that the EXPO model is superior to the IRT model when examining the trajectories of CNCCs at the early stage. As shown in the top panel in Figure 2, in Guangdong, China, model R^2 = 1.0, 0.99 and MAPE = 0.02, 0.06 for the EXPO and IRT models, respectively. In comparison to New York (US), model R^2 = 1.0, 1.0 and MAPE = 0.01, 0.29 for the EXPO and IRT models, respectively.

The following figure shows a comparison between R^2 (at the earlier 20-day stage) and MAPE (at the following 7-day stage). We note that there is no significant difference in R2 or MAPE between these 2 models when measuring ImpactCOVID on continents at the early outbreak stage.

3.1.2. Comparison of Angle Indexes at the Early Outbreak Stage. In Figure 4, we present the angle indexes based on the early outbreak stage of 20 days. Higher angle indexes are observed in Hubei Province (including Wuhan) in China as well as in New York and New Jersey in the United States. Darker colors indicate countries severely affected by COVID-19, including Iran and Turkey. The angle indexes were calculated using Eq. 1 below. The reader is encouraged to scan the QR code in Figure 4 to examine the Ogive curves for countries/regions when the color region is clicked.

3.2. Task 2

3.2.1. Comparison of the EXPO and IRT models for 2020. Modeling the CNCC data in the EXPO and IRT models to estimate their parameters in the long run revealed that the IRT model has significantly higher R^2 and smaller MAPE values, as shown in Figure 5. Forest plots were generated by computing the respective pair statistics of the mean and SD values in R^2 and MAPE, as shown in the upper and lower panels, respectively.

Currently, only South America and the United States have the same R^2 (>0.90) between the 2 models; see the upper panel of Figure 5. In other continents and in China, the IRT model has a higher R^2 and a smaller MAPE than the EXPO model;
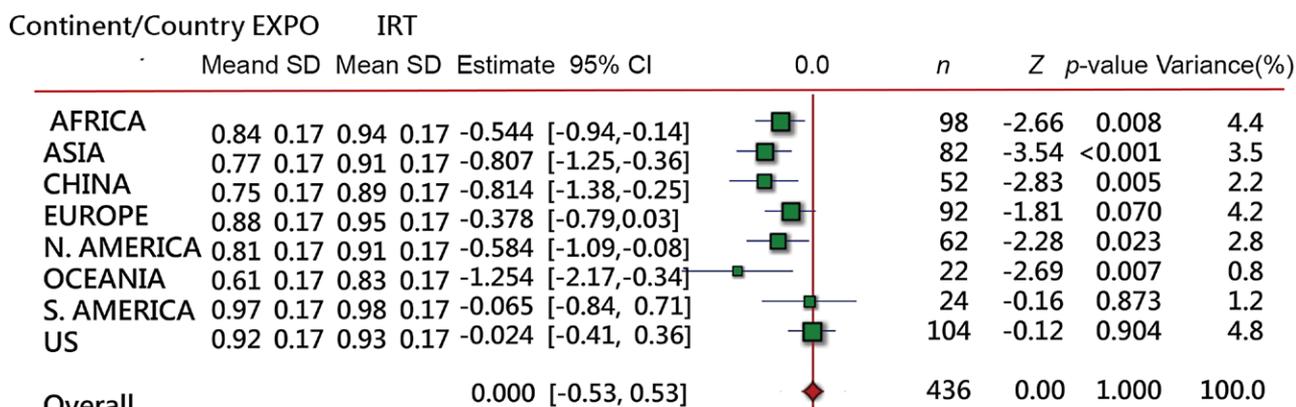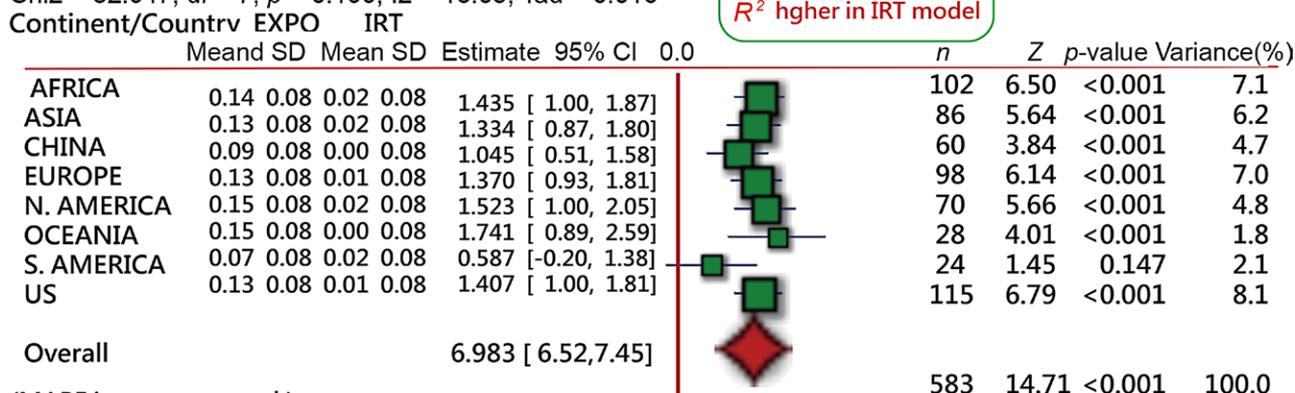
Figure 5. Comparison of the R² and MAPE values between the EXPO and IRT models using forest plots. EXPO = exponential growth, IRT = item response theory, MAPE = mean absolute percentage error.

see Figure 5 for a higher R² in the left column (which favors the IRT) and a smaller MAPE in the right column (which also favors the IRT).

3.2.2. Comparison of angle indexes in the long run. Figure 6 depicts the angle indexes based on the 2020 data. The most negatively impacted countries were India, the United Kingdom, and the United States (specifically California). In Figure 7, the angle is computed by the ACNCC (=CNCC at IP * CNCC at IP-6) divided by the LID (=6). In this regard, ImpactCOVID reflects the meaning of the disease outbreak since a 7-day incubation period is used to compute the angle index.

The reader is also invited to scan the QR code in Figure 6 to view the ogive curves for countries/regions of interest when the color region is clicked. For example, the EXPO and IRT models are used to compare the 2 ogive curves of

Hubei Province (including Wuhan) in China. Even though the 2 curves appear similar, the projection curve in the IRT model (shown at the bottom of Fig. 6) has a higher R² and a smaller MAPE than the EXPO model (0.99 vs 0.98 and 0% vs 0.2%, respectively).

3.3. Task 3

3.3.1. IP days in these 2 models on vaccination uptake per 100 people. In Figure 7, line-chart plots were used to compare vaccination rates per 100 people. Based on Figure 8, it can be seen that the IRT model performed better than the EXPO model for the CNCCs with a given Ogive curve (e.g., China and Algeria) and different IP days resulted from the 2 models (Fig. 8A), which significantly influenced the computation of ImpactCOVID using the angle index.
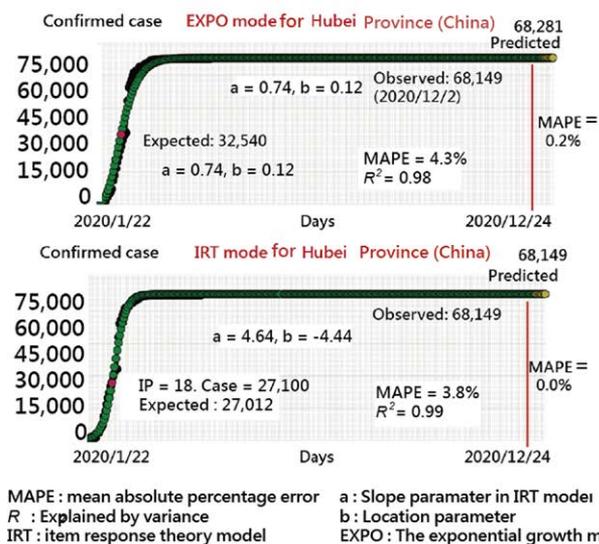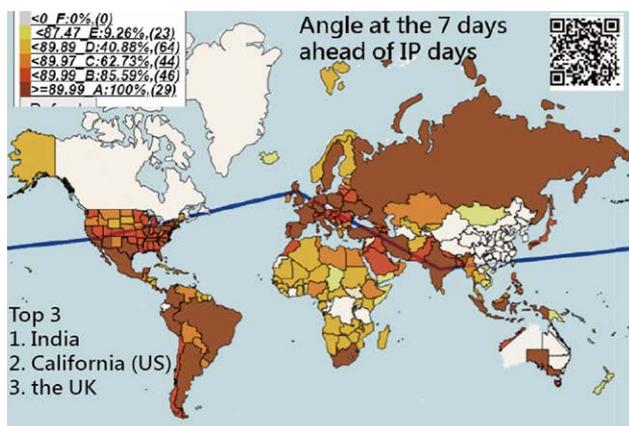
Figure 6. Top 3 countries hit by ImpactCOVID based on the angle index (top) and the comparison of modeling effects between the EXPO and IRT models (bottom). EXPO = exponential growth, ImpactCOVID = impacts of the COVID-19 outbreak, IRT = item response theory.

3.3.2. Comparison of IP days using the IRT model.
Figure 8B shows that Europe has the shortest IP days, followed by the United States and South America. In contrast, Africa had longer IP days in 2021.

3.4. Task 4

3.4.1. Comparison of goodness-of-fit measures.
We compared the 5 goodness-of-fit measures in terms of the usefulness of the model. Only the AUC possesses the feature that a high score indicates a more appropriate model for the COVID-19 pandemic. According to Figure 9, the EXPO model is more appropriate for early outbreak stages, while the IRT model is more appropriate for the entire year. Note that all paired data obtained from the 2 models were standardized into $\sim N(0,1)$ due to significantly wider ranges among the 5 indicators. Thus, the SMD can be compared using the forest plot in Figure 9.

3.4.2. Comparison of AUC indicators.
According to the scatter plot with the 95% control lines (Fig. 10), (1) the US states have higher AUCs (i.e., the determination coefficient in regression analysis) than the provinces/metropolitan cities in China. Most countries/regions lie along the 45-degree identity line. The AUC in Ukraine, for example, is 0.99, and in Taiwan,

it is 0.96. The AUC of other countries can be examined by scanning the QR code and clicking on the bubble of interest on the dashboard we designed on Google Maps.

3.5. Online dashboards shown on Google Maps
The QR codes in the figures are linked to the dashboards. Readers are recommended to view the dashboards on Google Maps.

4. Discussion

4.1. Principal findings
We observed that (1) EXPO and IRT models had identical PP at the earlier outbreak stage; (2) the IRT model had a higher model R^2 and smaller MAPE than the EXPO model in 2020; (3) Hubei Province in China had the highest angle index at the early stage; and (4) India, California (US), and the United Kingdom had the highest angle indexes in 2020. The IRT model is superior to the EXPO model when determining the IP on an Ogive curve.

4.2. Contributions of the study
Although the IP case index has been applied to examine the effective control of COVID-19,[2,20] the angle index can accurately evaluate ImpactCOVID, as shown in Figures 4 and 6.

Many researchers[10–21] have proposed the use of mathematical models to predict the number of COVID-19 cases, and others have investigated IP days during the COVID-19 pandemic.[32–36] However, no one used the IP days to compare the ImpactCOVID or applied the angle index to inspect the ImpactCOVID in countries/regions.

The second contribution of this study is the comparison of mathematical COVID-19 models using forest plots. To date, only a few studies have compared the MA and PP between models because of a lack of familiarity with the algorithms proposed by other authors. In this work, we compared 2 mathematical models[2,18] based on common conditions (e.g., the evolution of CNCC and ΔIP = 7 across countries/regions).

The EXPO model[18] has been verified in several regions in China, including Wuhan in Hubei Province, Guangdong Province, and other parts of mainland China, during the early outbreak stage (from January 27 to February 18, 2020). For comparison, this study featured (1) 2 time-frame stages (i.e., the early 20-day outbreak stage and the entire year of 2020) and (2) all countries/regions hit by ImpactCOVID using the angle index.

The quadratic equation model (QE) model[19] has been used to present the projected cases in Colombia and the deaths in Russia, India, and the rest of the world using past 31-day data up to May 29, 2020. In Tsai et al,[37] the authors set the constrained term at the middle point (i.e., $P(x2, y2)$) of the observations with exponential growth during the COVID-19 epidemic and found that the IRT model is superior to the QE model.

Using less constrained parameters makes the model a good fit for the data.[36,38] However, the one constrained term set at the middle point (i.e., $P(x2, y2)$) of the observations in the QE model[19] yields a low MA, which has been verified in a previous study.[37]

Two other studies[20,21] applied the IRT[39,40] to construct an Ogive curve and determine the IP days used for predicting the projected cases in a country/region based on the CNCCs. Nonetheless, a sophisticated analysis to determine the IP days using the Newton–Raphson Iteration Method[21,41–43] must be conducted in future studies.

4.3. Implications and recommendations
Numerous mathematical COVID-19 models[10–21] and IP determinations[2,8,9] have been proposed. However, none developed
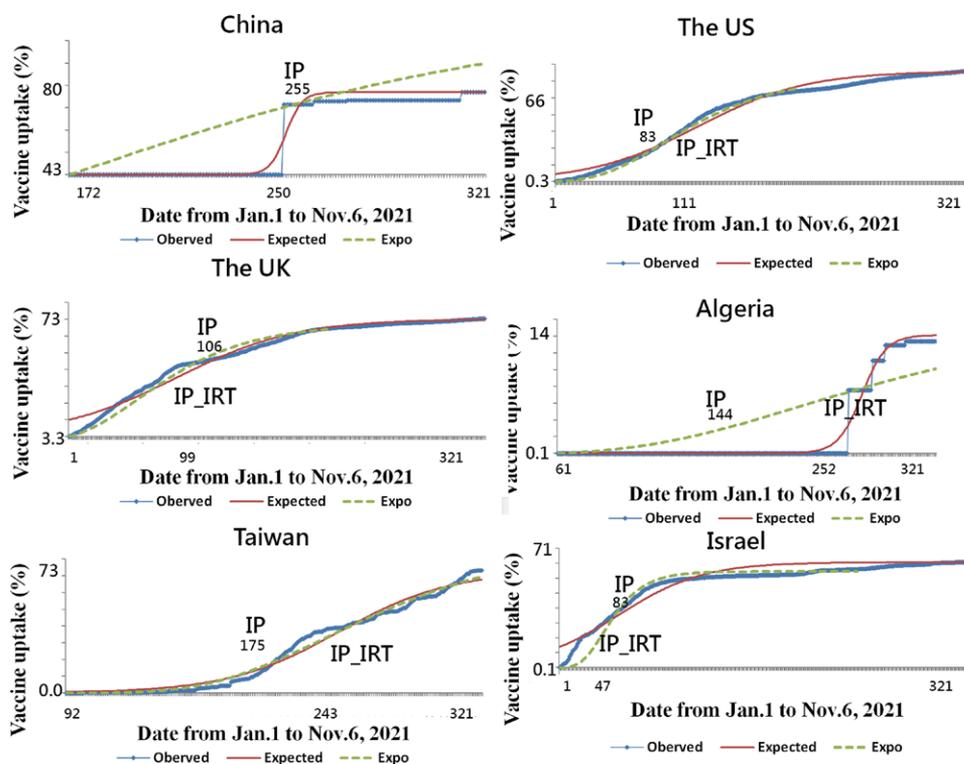
Figure 7. Comparisons of IP day yields by the 2 models using line-chart plots. IP = inflection point.

an angle index that can be used to measure ImpactCOVID and overcome the problem of the IP case index.[2,20]

Many online real- or near-real-time dashboards have been launched to track the worldwide spread of the COVID-19 outbreak.[44–49] However, most of these tools are similar to other traditional websites[50–54] and merely provide the same information as the WHO COVID-Dashboard.[45] An accurate assessment of COVID-19 requires further mathematical analyses of global data.[49] Although dashboards (e.g., JHU,[43] WHO,[45] and others[44,46,47]) have provided interesting visualizations for reporting the current state of COVID-19, these presentations lack important information on the disease outbreak using mathematical models to predict the projection of CNCCs in the future and understand COVID-19 trends[19] further in data.

One study[53] assessed 158 public Web-based COVID-19 dashboards and found that only a few dashboards employed predictive analytics by simulating various future scenarios. Thus, the imprecise predictive models and simulations early in the pandemic may have restricted their application. The current study presented predictive approaches with an online dashboard design that can benefit policymakers and epidemiologists during the COVID-19 epidemic.

4.4. Strengths of the study

First, the comparison of ImpactCOVID using the angle index in countries/regions can be applied to future relevant studies and is not limited to those merely focusing on the COVID-19 pandemic.

Second, MP4 videos on how to model the CNCCs and estimate the parameters for mathematical models[23] have been provided to ordinary readers who are familiar with Microsoft Excel and hope to replicate the study in the future.

Third, the use of Microsoft Solver add-in to estimate the model parameters is a common approach that can be easily adopted by researchers.[2,4,54–56] Data and model-building videos are provided in Supplemental Digital Contents 1, http://links.lww.com/MD/G983 and 2, http://links.lww.com/MD/

G984. Approaches for the search of IP days that can also be used for computing the angle index have been previously studied.[2,20,21]

Fourth, the choropleth and forest plots used in this study can provide comprehensive insights into the difference in the comparison of various countries/regions or pair-panel statistics. In turn, these differences can be used by policymakers and decision-makers in visualizing their data.

Furthermore, an MS Excel module for drawing the forest plot is provided. Readers are recommended to watch the abstract video and the Excel module presented in Chien.[23]

4.5. Limitations and future studies

Our study has several limitations. First, only 2 models were compared for MA and PP assessment. Future investigations are required to study other COVID-19 models and further understand their differences and merits.

Second, only short- and long-term time periods during the COVID-19 pandemic were compared using the forest plot in the 2 models. Whether differences in MA and PP exist in the medium-term (or mid-term) epidemic must be investigated in the future.

Third, the case number is changeable and may vary day by day, particularly in countries undergoing second or third waves (peaks) in the ongoing pandemic. Thus, model parameters and angle indexes would also vary with time.

Fourth, the Microsoft Solver add-in is not unique in estimating model parameters. Many other methods and mathematical techniques should be used in making estimations and comparisons in the future, such as the Newton–Raphson iteration method[21,41–43] for searching IPs on a given ogive curve.[21]

Fifth, visual dashboards on Google Maps are not free of charge, and a paid project key for using the Google cloud platform is needed. One limitation in using the dashboard is that it cannot be easily replicated by other authors or programmers for use in a short period of time.

A   Comparison of IPs between models

Continent	EXPO model Mean	SD	IRT model Mean	SD	SMD 0.0	Measure	95% CI	n	Z	p-value	weight(%)
1. Africa	186.08	44.10	240.53	56.57		1.066	(-1.47, -0.66)	53	-5.17	<0.001	23.8
2. Asia	155.78	45.44	214.52	60.15		1.094	(-1.51, -0.68)	50	-5.13	<0.001	22.3
3. Europe	111.10	33.02	159.46	42.68		1.258	(-1.68, -0.83)	50	-5.78	<0.001	21.4
4. North America	155.97	52.51	184.82	59.64		0.508	(-0.98, -0.03)	34	-2.08	0.038	17.1
5. Oceania	207.82	76.55	216.53	68.31		0.117	(-0.77, 0.54)	17	-0.35	0.726	9.0
6. South America	138.00	46.71	190.23	70.53		0.846	(-1.62, -0.07)	13	-2.13	0.033	6.4
Overall						0.918	(-1.12, -0.72)	217	-9.12	<0.001	100.0

Q=12.207, df= 5, p=0.032, I2=59.04, Tau=0.091

Favors IRT   Favors EXPO

B   Comparison of IPs determininded by IRT model for continents

Continent	IRT model Mean	SE	SMD 0.0	Measure	95% CI	n	Z	p-value	weight(%)
1. Africa	240.53	7.84		1.382	(1.83, 0.94)	53	-6.07	<0.001	8.3
2. Asia	214.52	8.59		0.472	(0.83, 0.11)	50	-2.57	0.010	12.7
3. Europe	159.46	6.10		-1.453	(-0.81,-2.10)	50	4.40	<0.001	3.9
4. North America	184.82	10.38		-0.566	(-0.41,-0.72)	34	7.18	<0.001	69.0
5. Oceania	216.53	17.08		0.542	(-0.07, 1.16)	17	1.73	0.084	4.4
6. South America	190.23	20.36		-0.377	(-1.37, 0.62)	13	-0.74	0.459	1.7
Overall				-0.256	(-0.38,-0.13)	217	-1.48	0.140	100.0

Q=102.665, df= 5, p<0.001
I2=95.13, Tau=0.844

Favors IP shorter   Favors IP longer

Figure 8. Comparisons of IP day yields by the 2 models across continents and IP comparison among continents. IP = inflection point.

Sixth, only 5 goodness-of-fit measures were applied to compare model usefulness (Fig. 10). Other indicators (e.g., Hannan–Quinn criterion[57] and minimum description length[58]) could be employed in future studies.

Although IRT has superior MA and PP in the long run, other user-friendly mathematical models are also available for readers to understand the properties of exponential growth[18] and quadratic equations[19] in nature and for scientists to build improved COVID-19 models for further comparisons.

5. Conclusions

The 2 proposed models were compared according to their MA, PP, and 5 goodness-fit measures in 2 time-frame scenarios,

Type	Inducator	EXPO model Mean	SD	IRT model Standardized score with indictor Mean	SE	SMD 0.0	Measure	95% CI	n	Z	p-value	weight (%)
Early 20 days	1. AUC●	-0.14	1.14	-0.10	1.47		-0.03	(-0.21 - 0.15)	230	-0.34	0.734	10.4
	2. residuals	-0.06	0.00	0.06	1.38		-0.13	(-0.31 - 0.06)	230	-1.34	0.180	10.4
	3. AIC	-0.59	0.75	0.48	0.96		-1.24	(-1.44 --1.04)	230	-12.19	<0.001	8.7
	4. BIC	-0.50	1.07	0.18	1.26		-0.58	(-0.76 --0.39)	230	-6.09	<0.001	10.0
	5. Brier score	-0.59	0.75	0.48	0.96		-1.24	(-1.44 --1.04)	230	-12.19	<0.001	8.7
Entire year 2020	6. AUC●	-0.99	2.33	-0.08	1.24		-0.49	(-0.68 --0.31)	230	-5.20	<0.001	10.1
	7. residuals	0.07	1.42	-0.07	0.00		0.15	(-0.04 - 0.33)	230	1.57	0.116	10.4
	8. AIC	0.01	1.42	-0.02	0.14		0.04	(-0.15 - 0.22)	230	0.39	0.697	10.4
	9. BIC	0.07	1.44	-0.29	1.23		0.27	(0.09 - 0.45)	230	2.90	0.004	10.3
	10.Brier score	0.01	1.42	-0.02	0.14		0.04	(-0.15 - 0.22)	230	0.39	0.697	10.4
Overall							-0.29	(-0.35 - -0.23)	230	-9.65	<0.001	100.0

Q=281.49 df=9 P<.001
I2=96.8 Tau=0.274

● Favors IRT
● Favors EXPO
● the larger the better

Larger in IRT   Larger in EXPO

Figure 9. Comparisons of goodness-of-fit measures for models in 2 scenarios.
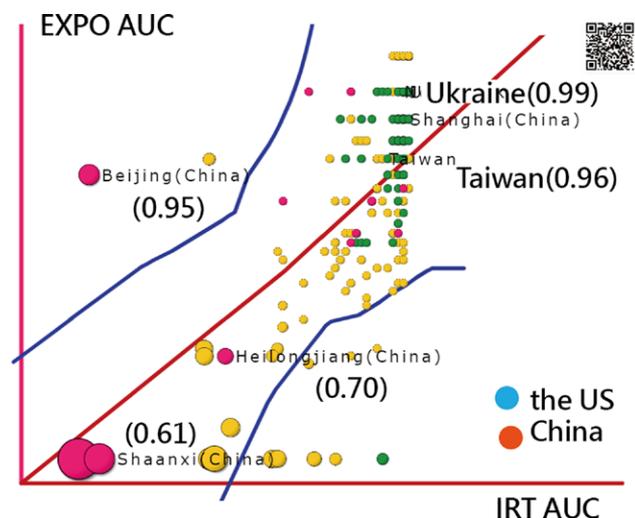
Figure 10. Comparisons of AUC yielded by the 2 models for countries/regions on the scatter plot with the 95% control lines. AUC = area under the curve.

namely, the early outbreak stage and throughout the entire year of 2020. We found that the IRT model is superior to the EXPO model in the long term. However, the EXPO model is better than the IRT model at the early stage because only the Ogive curves in the IRT model can appropriately generate the IP, which can then be used to compute the angle index. Both models are recommended so that readers can accurately project the CNCCs in other outbreak scenarios in the future. In addition, these models are not limited to the COVID-19 pandemic.

Author contributions

T.-W.C. developed the study. H.Y., K.T., C.Y., and W.C. analyzed the data. K.T. monitored the process of this study and helped respond to the reviewers' advice and comments. T.-W.C. drafted the manuscript, and all authors provided critical revisions for important intellectual content. The study was supervised by K.T. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

References

[1] Wiki. Disease outbreak. Available at: https://en.wikipedia.org/wiki/Disease_outbreak [access date July 24, 2022].

[2] Wang L-Y, Chien T-W, Chou W. Using the ipcase index with inflection points and the corresponding case numbers to identify the impact hit by COVID-19 in China: an observation study. Int J Environ Res Public Health. 2021;18:1994.

[3] Chang CS, Yeh YT, Chien TW, et al. The computation of case fatality rate for novel coronavirus (COVID-19) based on Bayes theorem: an observational study. Medicine. 2020;99:e19925.

[4] Fisman D, Rivers C, Lofgren E, et al. Estimation of MERS-coronavirus reproductive number and case fatality rate for the spring 2014 Saudi Arabia outbreak: insights from publicly available data. PLoS Curr. 2014;6:6.

[5] Yie K-Y, Chien T-W, Yeh Y-T, et al. Using social network analysis to identify spatiotemporal spread patterns of COVID-19 around the world: online dashboard development. Int J Environ Res Public Health. 2021;18:2461.

[6] Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci USA. 2005;102:16569–72.

[7] Fenner T, Harris M, Levene M, et al. A novel bibliometric index with a simple geometric interpretation. PLoS One. 2018;13:e0200098.

[8] Chien, T.W. Trend of COVID-19 using 7-day mean number of confirmed cases. Available at: http://www.healthup.org.tw/html100/searchpubmed.asp [Access date March 10, 2021].

[9] Fan RG, Wang YB, Luo M, et al. SEIR-Based COVID-19 transmission model, and inflection point prediction analysis. J Univ Electron Sci Technol China. 2020;49:369–74.

[10] Perc M, Miksić NG, Slavinec M, et al. Forecasting COVID-19. Front Phys. 2020;8:127.

[11] Fang Y, Nie Y, Penny M. Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: a data-driven analysis. J Med Virol. 2020;92:645–59.

[12] Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modeling study. Lancet. 2020;395:689–97.

[13] Anastassopoulou C, Russo L, Tsakris A, et al. Data-based analysis, modeling and forecasting of the COVID-19 out-break. PLoS One. 2020;15:e0230405.

[14] Zhao S, Chen H. Modeling the epidemic dynamics and control of COVID-19 outbreak in China. Quant Biol. 2020;8:11–9.

[15] Rong X, Yang L, Chu H, et al. Effect of delay in diagnosis on transmission of COVID-19. Math Biosci Eng. 2020;17:2725–40.

[16] Mandal M, Jana S, Nandi SK, et al. model based study on the dynamics of COVID-19: prediction and control. Chaos Solitons Fractals. 2020;136:109889.

[17] Huang J, Qi G. Effects of control measures on the dynamics of COVID-19 and double-peak behavior in Spain. Nonlinear Dyn. 2020;101:1889–99.

[18] Shang C, Yang Y, Chen G-Y, et al. simple transmission dynamics model for predicting the evolution of COVID-19 under control measures in China. Epidemiol Infect. 2021;149:1–10.

[19] Florez H, Singh S. Online dashboard and data analysis approach for assessing COVID-19 case and death data. F1000 Res. 2020;9:570.

[20] Lee KW, Chien TW, Yeh YT, et al. An online time-to-event dashboard comparing the effective control of COVID-19 among continents using the inflection point on an Ogive curve: observational study. Medicine (Baltim). 2021;100:e24749.

[21] Ho SY, Chien TW, Shao Y, et al. Visualizing the features of inflection point shown on a temporal bar graph using the data of COVID-19 pandemic. Medicine (Baltim). 2022;101:e28749.

[22] Google Team. 2019 Novel coronavirus (nCoV) data repository. Available at: https://github.com/CSSEGISandData/2019-nCoV [Access date October 19, 2020].

[23] Chien TW. Data and MP4 video of model building. Available at https://osf.io/4pmn7/?view_only=5afba393509746bd8a332b2c65839034 [Access date March 19, 2022].

[24] Shang X-D, Tong P, Xia K-Q. Scaling of the local convective heat flux in turbulent Rayleigh-Bénard convection. Phys Rev Lett. 2008;100:244503.

[25] Shang X-D, Liang C-R, Chen G-Y. Spatial distribution of turbulent mixing in the upper ocean of the South China Sea. Ocean Sci. 2017;13:503–19.

[26] Yan Y-H, Chien T-W. The use of forest plot to identify article similarity and differences in characteristics between journals using medical subject headings terms. Medicine. 2021;100:e24610.

[27] Chien T-W, Wang H-Y, Hsu C-F, et al. Choropleth map legend design for visualizing the most influential areas in article citation disparities. Medicine. 2019;98:e17527.

[28] Wright BD, Stone MH. Best Test Design (Rasch Measurement Series). Chicago, IL 60637: MESA Press, 5835 S. Kimbark Avenue. 1979:p94.

[29] Akaike H. Information theory and extension of the maximum likelihood principle. In: Kotz S, Johnson NL (eds), Breakthroughs in Statistics Volume 1. Foundations and Basic Theory. New York, NY: Springer-Verlag. 1973:610–24.

[30] Schwarz G. Estimating the dimension of a model. Ann Stat. 1979;6:461–4.

[31] Brier GW. Verification of forecasts expressed in terms of probability. Mon Weather Rev. 1950;78:1–3.

[32] Buan Q, Wu J, Wu G, et al. Predication of inflection point and outbreak size of COVID-19 in new epicenters. Nonlinear Dyn. 2020;101:1561–81.

[33] Chatham WW. Treating Covid-19 at the inflection point. J Rheumatol. 2020;47:1–10.

[34] Gu C, Zhu J, Sun Y, et al. The inflection point about COVID-19 may have passed. Sci Bull. 2020;65:865–7.

[35] Jeong GH, Lee HJ, Lee J, et al. Effective Control of COVID-19 in South Korea: cross-sectional study of epidemiological data. J Med Internet Res. 2020;22:e22103.

[36] Gallacher D, Kimani P, Stallard N. Extrapolating parametric survival models in health technology assessment: a si-mu-lation study. Med Decis Making. 2021;41:37–50.

[37] Tsai KT, Chien TW, Lin JK, et al. Comparison of prediction accuracies between mathematical models to make projections of confirmed cases during the COVID-19 pandemic by country/region. Medicine (Baltim). 2021;100:e28134.

[38] Van Rijn PW, Sinharay S, Haberman SJ, et al. Assessment of fit of item response theory models used in large-scale educational survey assessments. Large-Scale Assess Educ. 2016;4:1.

[39] Lord, F.M. Applications of Item Response Theory to Practical Testing Problems. Hillsdale, NJ: Lawrence Erlbaum Associates; 1980.

[40] NBSP; Johns Hopkins University Coronavirus Resource Center. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). Available at: https://coronavirus.jhu.edu/map.html [Access date October 29, 2020].

[41] Shao Y, Chien TW. The determination of inflection point on a given Ogive curve using the item response theory (IRT) model. J Bibliographical Anal Stat. 2021;18:31–3.

[42] Linacre JM. Rasch estimation: iteration and convergence. Rasch Measurement Transactions. 1987;1:7–8

[43] Shao Y, Nadkarni S, Niu K, et al. Understanding of the Newton–Raphson iteration method in Rasch model. J Bibliographical Anal Stat. 2021;18:71–6.

[44] Leszkiewicz, A. Dashboard online for COVID-19 in near real time. Available at: https://avatorl.org/covid-19/ [Access date February 20, 2021].

[45] World Health Organization. WHO coronavirus disease (COVID-19) dashboard. 2020. Available at: https://covid19.who.int/ [Access date March 3, 2021].

[46] HealthMap. Novel coronavirus 2019-nCoV (Interactive Map). Available at: https://healthmap.org/wuhan/ [Access date February 20, 2021].

[47] Schiffmann, A. A creator of one COVID-19 dashboard. Available at: https://ncov2019.live/data [Access date February 20, 2021].

[48] Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis. 2020;20:533–4.

[49] WHO. Novel coronavirus (2019-nCoV) outbreak. Available at: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports [Access date August 10, 2020].

[50] Centers for Disease Control and Prevention. Available at: http://www.cdc.gov/ [Access date February 6, 2013].

[51] European Centre for Disease Prevention and Control (ECDC). Novel coronavirus. Available at: https://www.ecdc.europa.eu/en/home [Access date February 20, 2021].

[52] DXY. Novel coronavirus (2019-nCoV). Available at: https://ncov.dxy.cn/ncovh5/view/pneumonia. [Access date February 20, 2021].

[53] Ivanković D, Barbazza E, Bos V, et al. Features constituting actionable COVID-19 dashboards: descriptive assessment and expert appraisal of 158 public web-based COVID-19 dashboards. J Med Internet Res. 2021;23:e25682.

[54] Yan Y-H, Chien T-W, Yeh Y-T, et al. An app for classifying personal mental illness at workplace using fit statistics and convolutional neural networks: survey-based quantitative study. JMIR mHealth uHealth. 2020;8:e17857.

[55] Ma S-C, Chou W, Chien T-W, et al. An app for detecting bullying of nurses using convolutional neural networks and web-based computerized adaptive testing: development and usability study. JMIR mHealth uHealth. 2020;8:e16747.

[56] Lee Y-L, Chou W, Chien T-W, et al. An app developed for detecting nurse burnouts using the convolutional neural networks in microsoft excel: population-based questionnaire study. JMIR Med Inform. 2020;8:e16528.

[57] Burnham KP, Anderson DR. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. 2nd ed. New York Berlin Heidelberg, US: Springer-Verlag; 2002:P287. Available at: https://caestuaries.opennrm.org/assets/06942155460a79991fdf1b-57f641b1b4/application/pdf/burnham_anderson2002.pdf [access date July 24, 2022].

[58] Rissanen J. Modeling by shortest data description. Automatica. 1978;14:465–71.